

firestar—advances in the prediction of functionally important residues

Gonzalo Lopez^{1,2}, Paolo Maietta¹, Jose Manuel Rodriguez¹, Alfonso Valencia¹ and Michael L. Tress*

¹Structural Computational Biology Group, Spanish National Cancer Research Centre (CNIO), c. Melchor Fernandez Almagro, 3, 28029 Madrid, Spain and ²Center for Computational Biology and Bioinformatics, Columbia University, New York, NY 10032, USA

Received February 17, 2011; Revised May 12, 2011; Accepted May 13, 2011

ABSTRACT

firestar is a server for predicting catalytic and ligand-binding residues in protein sequences. Here, we present the important developments since the first release of **firestar**. Previous versions of the server required human interpretation of the results; the server is now fully automatized. **firestar** has been implemented as a web service and can now be run in high-throughput mode. Prediction coverage has been greatly improved with the extension of the FireDB database and the addition of alignments generated by HHsearch. Ligands in FireDB are now classified for biological relevance. Many of the changes have been motivated by the critical assessment of techniques for protein structure prediction (CASP) ligand-binding prediction experiment, which provided us with a framework to test the performance of **firestar**. URL: <http://firedb.bioinfo.cnio.es/Php/FireStar.php>.

INTRODUCTION

The ultimate goal for researchers working with experimental protein sequences is to determine function. Computational methods form the basis of initial approaches to function determination because most approaches for the characterization of molecular function are difficult, expensive and time consuming. Many methods have been developed to predict protein function in recent years and the power of homology-based function prediction methods has increased thanks to the prodigious growth in the sequence and structural databases that are due to genome sequencing projects (1) and structural genomics initiatives (2) have increased the power of homology-based function prediction methods.

As the structural databases expand and populate structural space, a great deal of interesting biological

information is being generated. Much of this, such as the amino acid residues implicated in molecular interactions or catalysis, can be found at the residue level.

FireDB (3) and the **firestar** web server (4) were developed specifically to make use of this data in order to predict biologically important residues in protein sequences. FireDB is a database of annotated catalytic residues and ligand-binding residues culled from the protein structures deposited in the Protein Data Bank (PDB, 5). **firestar** uses the functional information in FireDB to make predictions of ligand-binding residues and catalytic residues.

The identification of potential ligand-binding or catalytic residues can provide important clues for the design of targeted biochemical experiments, and can be a vital part of drug design and virtual screening. Ligand-binding site predictions can also be helpful in predicting general protein function, while predicted binding sites may also act as anchoring regions in the generation of structural models. Baker *et al.* (6) used predicted zinc-binding residues as an important constraint to limit the structural space of possible decoys in their ROSETTA algorithm.

A number of ligand-binding prediction methods have been published since 2007 (7,8), mostly motivated by the critical assessment of techniques for protein structure prediction (CASP) ligand-binding prediction experiments (9,10), which provided a blind test framework for the evaluation of ligand-binding methods.

Here, we present the new developments of **firestar**. Several new features have been incorporated into the server to improve the quality of the predictions and the usability of the web interface. CASP blind tests show that **firestar** predictions are state of the art.

DESCRIPTION OF THE TOOL

We developed **firestar** with the aim of predicting functional residues from the information extracted from remotely related structures. The server makes predictions

*To whom correspondence should be addressed. Tel: 34 91 732 8000; Fax: +34 91 224 6980; Email: mtress@cnio.es

based on local sequence conservation matches to the biologically relevant small molecule ligand binding residues in FireDB and annotated catalytic residues from the Catalytic Site Atlas (CSA, 11).

Protocol

The *firestar* web server works as follows:

- (1) Most users will input a single protein sequence, but there is also an option to search with a protein structure, either directly from the PDB or user uploaded. The sequence is extracted from the 3D structure.
- (2) PSI-BLAST (12) profiles are generated for the sequences from a locally generated 70% redundant database. The profiles are used to search against the FireDB template database.
- (3) Users may specify the BLAST *e*-value cut-off for the final search of the FireDB templates; note that the default *e*-value is intentionally high as functional information will be present in even very distantly related proteins.
- (4) At the same time HHsearch (13) uses hidden Markov models generated from PSI-BLAST sequences to search against a profile database created *ad hoc* from all the FireDB template sequences. *firestar* uses all the templates detected by HHsearch to predict binding residues.
- (5) Both sets of alignments between query sequences and FireDB templates, with their accompanying functional information, are used to predict functional sites and likely bound ligands.
- (6) The predicted sites are evaluated by SQUARE (14).
- (7) The combined results from the HHsearch and PSI-BLAST searches are displayed on the main output page and the predicted functional residues are highlighted (example output shown, see Figure 1).
- (8) Users can also browse the alignments generated by the HHsearch and PSI-BLAST searches. Here, the local conservation score for each aligned pair of residues is shown in shades of blue, the darker the colour, the higher the local conservation score.
- (9) If the users submit a structure they can generate structural alignments with the FireDB templates using the LGA structural alignment method (15) and visualise the alignments using Jmol.

There is more detailed information in the online help pages.

EVALUATING *firestar* PERFORMANCE

firestar has been tested during the CASP7, CASP8 and CASP9 ligand-binding prediction experiments (9,10). The CASP experiments are the best testing ground for web servers, although results from the CASP ligand-binding prediction experiment should be taken with care—each CASP is a snapshot of the predictive capacity of servers and human groups over a limited time period and over a limited set of targets. Nevertheless, the results from the three CASP experiments

form a body of evidence, which suggests that *firestar* is a state of the art ligand-binding predictor.

The server was not allowed to participate officially in either the CASP7 or the CASP8 experiments because the authors were also CASP assessors. In CASP8, *firestar* made blind predictions during the prediction season under the same rules as other experimental groups, and we evaluated the *firestar* predictions along with the other servers. The CASP ligand-binding prediction experiments use Matthews correlation coefficients (MCC) to evaluate all predictions against the known ligand-binding residues. The MCC is a measure of binary classification quality. It combines true positives, true negatives, false positives and false negatives, and one advantage is that it can be used when the two classifications are of very different sizes, as they often are with binding and non-binding residues. MCC values are between -1 and +1, where 1 represents a perfect prediction and 0 a random prediction.

Over the CASP7 and CASP8 experiments, *firestar* correctly predicted the ligand-binding sites for the 46 targets that bound biologically relevant ligands and for which it made predictions. There were two targets with biologically relevant ligands for which *firestar* did not make a prediction. In CASP8, *firestar* obtained an MCC score of 0.754 over the 26 targets it predicted (see Supplementary Table S1). The sensitivity of the *firestar* predictions in CASP8 was 0.9 (90% of known functional residues were in the predictions), while the precision 0.67 (67% of predicted residues were known functional residues) suggesting a certain overprediction. Indeed, *firestar* was tuned to make predictions at a distance of 1.5 Å in CASP8, while the official distance used to define a ligand-binding residue was 0.5 Å. Most false positive predictions were residues that were next to the official binding site. Of 87 *firestar* false positives in CASP8, 63 were within 2.5 Å of the bound ligand. At a distance of 1.5 Å, *firestar* precision was 0.8 (80% of predicted residues were known functional residues), although the sensitivity dropped to 0.84. *firestar* had a better mean MCC score than all officially participating groups in CASP8, the human predictors as well as the server groups (Figures 1 and 2).

The *firestar* server participated in CASP9 and we can report the preliminary official results for the ligand-binding site prediction category. Over the 25 targets predicted by both servers, the I-TASSER server (16) had a similar performance to *firestar* (I-TASSER MCC of 0.7 and *firestar* MCC of 0.71). The human *firestar* and I-TASSER predictors were marginally better than their servers in head-to-head comparisons and two other human groups (neither of which have publicly available servers) had slightly lower MCC than *firestar*. The other 12 server groups that participated in CASP had substantially lower MCC scores than *firestar*.

Unfortunately, for technical reasons the results from CASP9 were not directly comparable with those of CASP8. CASP9 assessors had to include targets that bound to non-biological ligands such as solvents and buffers in the assessment. Over the small subset of 10 targets that did bind biological ligands, *firestar* had higher MCC scores than all server groups (a mean MCC

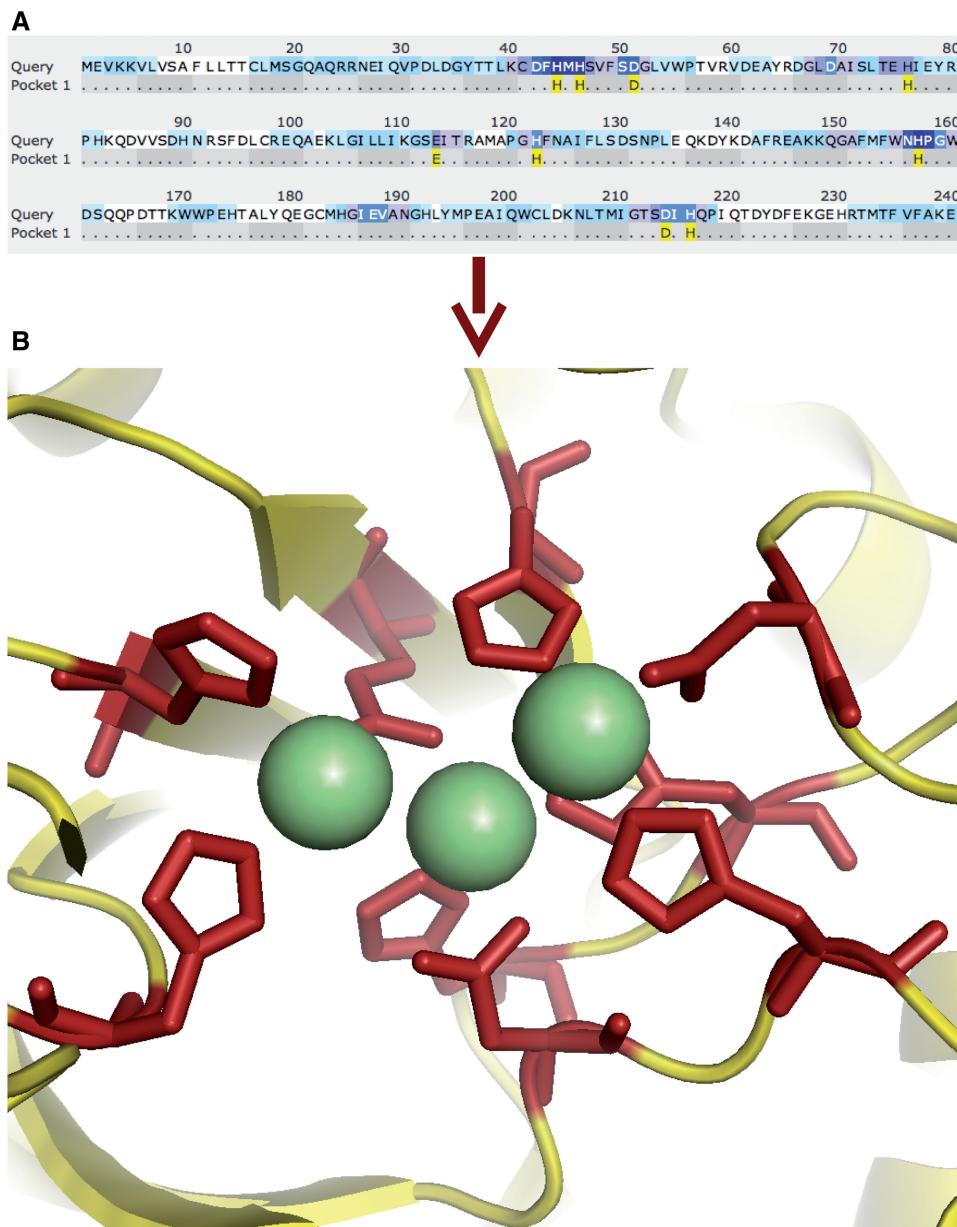


Figure 1. Outstanding *firestar* prediction for CASP8. The prediction for target T0407, 1 of 12 targets for which *firestar* would have recorded the best MCC score. (A) The prediction from *firestar*. The residues highlighted in yellow were the prediction. (B) T0407 was a predicted metal-dependent phosphoesterase and was crystallized with three calcium atoms (shown in light green). *firestar* predicted all the calcium-binding residues (shown in red) without any over prediction.

score of 0.72 against 0.65 for I-TASSER, see Supplementary Table S2).

In total, the automatic *firestar* server made predictions for 82 assessed targets over three CASP editions. The server failed to make a prediction for five targets (most because of a technical problem that has now been fixed). *firestar* correctly predicted the binding site for all targets, though not for all the binding residues. These predictions included the three free modelling targets (those without detectable structural templates) that bound ligands in the CASP7 and CASP8 editions. *firestar* was able to predict the binding sites for these targets because *firestar*

does not need to build accurate 3D models to make reliable predictions.

NEW ADDITIONS AND IMPROVEMENTS in *firestar*

The three main new developments in *firestar* have all contributed to huge improvements in the server. From a technical point of view, the server is now much easier to use and the fact that *firestar* now allows high-throughput predictions adds another dimension to the prediction of functional residues. The definition of the biological

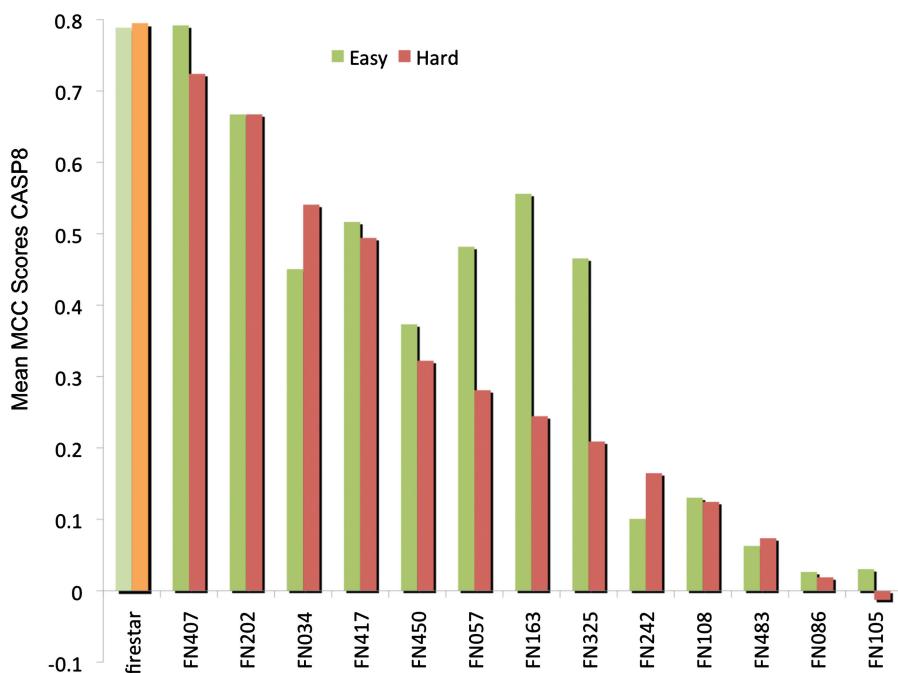


Figure 2. *firestar* performance during CASP8. Over the CASP8 targets, *firestar* obtained an MCC score of 0.75 when predicting residues in contact with ligands at a distance threshold of 0.5 Å plus van der Waals distances. The figure shows the targets separated into easier and harder targets based on their homology to known structures (10). *firestar* had higher MCC scores than all officially participating groups in CASP8, including human predictors.

relevance of bound ligands improves the accuracy of the predictions by removing false positives generated from ligands of dubious functional importance. The addition of HHsearch alignments increases the coverage of the *firestar* alignments.

Automatic interpretation

Predictions are made from the alignments generated with the templates in FireDB. Previous versions of *firestar* required human interpretation of the results—probable functional residues could be gleaned from a by-eye inspection of the pairwise alignments between the query sequence and the templates with bound ligands. The detailed results pages with the PSI-BLAST and HHsearch alignments are still available and are linked from the main output page. These extended results show all pairwise alignments between the query sequence and those FireDB templates that have functional annotations. These pages are important in those cases where the *firestar* summary pages do not return a result, because the alignments evaluated by SQUARE that are found in these pages can often give clues to possible binding sites.

Whereas the old version of *firestar* required expert input, the new process of predicting functional residues is completely automated. As previously the predictions from each PSI-BLAST and HHsearch alignment are evaluated separately, but now the predictions from each alignment are collated to generate an overall functional site prediction that is incorporated into a single results page. The graphical output (Figure 3A) shows the query amino acid chain coloured by relative local conservation

scores and highlights predicted catalytic (green) or ligand-binding (yellow) residues. Each pocket shown in this section is the result of merging predicted functional sites where at least 40% of residues overlap.

A text summary provides information for each individual predicted binding and catalytic site, including a list of predicted residues, the mean SQUARE score for the site and which ligands are found in the homologues. In the text summary, predicted binding sites with at least 60% residue overlap are merged. Sites that bind metals are differentiated from non-metal sites regardless of the overlapping percentage.

In addition to the summary page, there are two other levels of output available to the user, the detailed HHsearch and PSI-BLAST alignment evaluation pages (Figure 3B) and the raw PSI-BLAST/HHsearch output. The detailed alignment evaluation pages show the SQUARE evaluations of each template–target alignment and how these scores relate to ligand-binding and catalytic residues. The raw output contains all the target–template alignments, including those FireDB templates with no site information.

Alignments with HHsearch

The new *firestar* release includes HHsearch as an additional means of generating alignments between the query sequence and FireDB templates. HHsearch will find different homologues and (just as important) will create different alignments from PSI-BLAST. Both PSI-BLAST and HHsearch provide a pool of input alignments that are used to generate the initial prediction. Although the

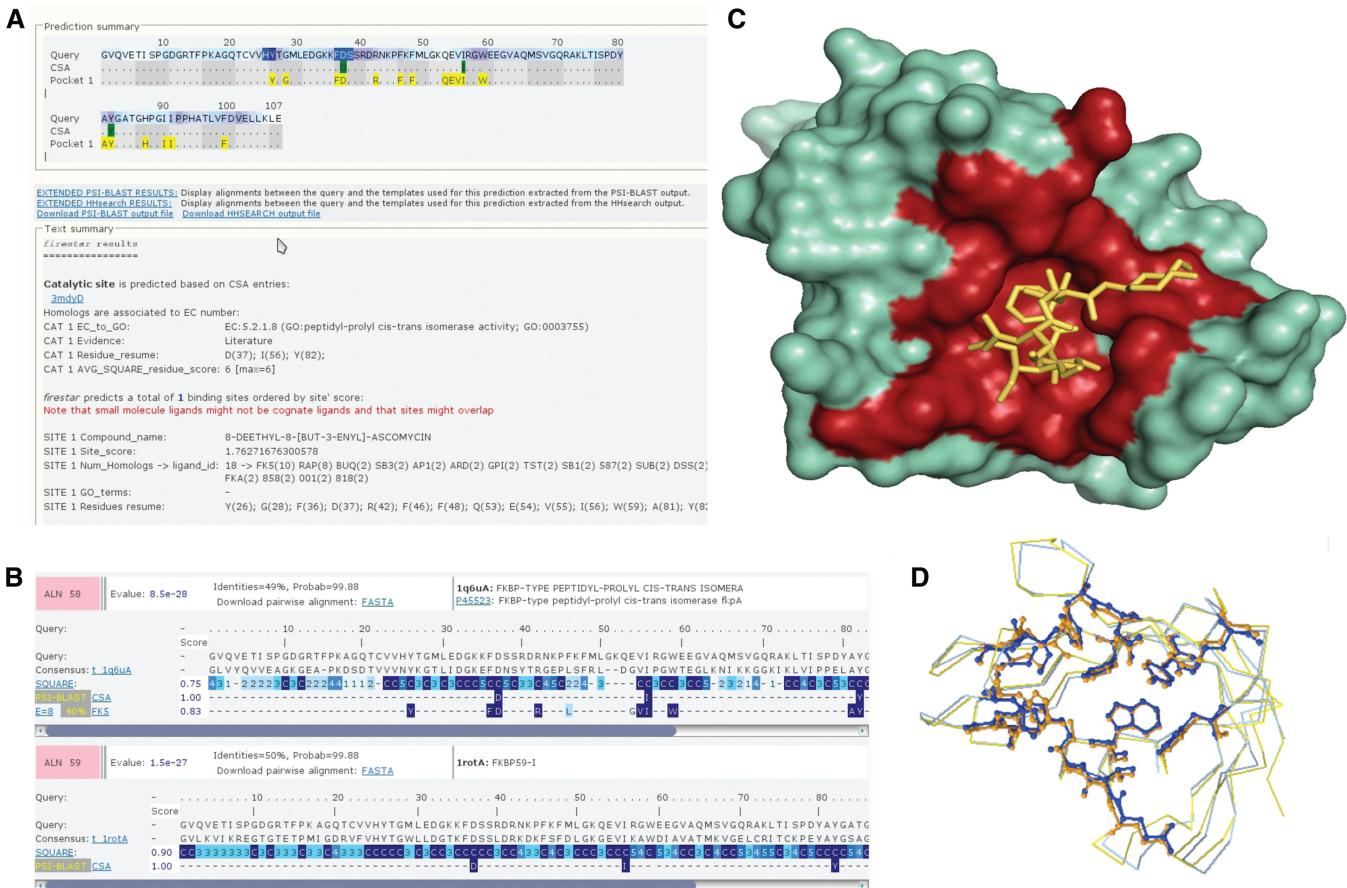


Figure 3. The new *firestar* interface. (A) Summary results page. In the upper part, the query amino acid sequence with predicted catalytic site residues (highlighted green residues) and binding site residues (yellow) shown on a single line. A text summary is displayed below for each prediction with a resume of the site score, the residues involved, and possible ligand if the site is ligand binding. (B) The HHsearch extended results page showing alignments between 1tcoC and two templates. The previous output style has been maintained, per-residue local conservation score is shown in blue (the darker the blue the more strong the local conservation) and the ligand-binding residues (or catalytic residues) in each FireDB template highlighted below the query-template alignment and the conservation score. (C) A PYMOL representation of the surface of PDB structure 1tcoC surface interacting with its inhibitor FK5 ('sticks'). The residues highlighted in red represent the *firestar* prediction from (A). (D) A Jmol representation of the LGA structural alignment between 1tcoC and the template 1q6uA. The Jmol applet integrated in *firestar* permits the visualization of the binding residues and/or catalytic residues ('sticks') of both structures.

alignments generated from HHsearch are in theory more powerful than PSI-BLAST (HHpred, based on HHsearch, was rated the best performing server in the official CASP9 evaluation at the meeting in December) the alignments from PSI-BLAST and HHsearch are complementary and equally valid for the prediction of ligand-binding residues in *firestar*.

Both methods are set up with lax cut-offs. The reason for this is that many of the short low-scoring local alignments generated by HHsearch and PSI-BLAST include functional information. With these cut-offs, the two methods will detect remote homologues, some false positive hits and many short alignments. However, this is not important because the alignments are only used as the initial input to *firestar*. The evaluation is carried out by SQUARE. SQUARE locates highly conserved local regions of residues (14) within the PSI-BLAST and HHsearch alignments through profile-profile comparison. Only those template ligand-binding residues that are in aligned regions with high local conservation (according

to SQUARE) can be considered as binding residues in the target. SQUARE has been shown to be particularly effective at predicting ligand-binding residues from alignments (17).

Once these potential binding residues are localized from all the alignments from HHsearch and PSI-BLAST, *firestar* determines whether the ligand-binding residues do form part of a functionally relevant binding site according to the numbers of residues detected (one limitation is that each binding site needs to be composed of a minimum number of highly conserved residues) and based on the biological relevance of the ligand.

The evaluation process weeds out the vast majority of the initial predictions. For example, HHsearch and PSI-BLAST generate 276 different alignments for the recent CASP target, T0614, but despite all the alignments no site was predicted for target T0614.

Above all, the effect of combining PSI-BLAST and HHsearch alignments is to extend the coverage of *firestar* predictions. The extended coverage will come

from two different sources: from those extra FireDB homologues that only HHsearch detects and from those alignments where HHsearch aligns correctly and PSI-BLAST does not.

For example, we have run *firestar* with HHsearch and PSI-BLAST alignments for part of the human genome. Adding HHsearch alignments increases coverage by 34%. We ran *firestar* for all 798 genes in chromosomes 21 and 22 annotated by Gencode in their 3C release (18). For the transcripts from these genes, PSI-BLAST–*firestar* predicts 12 657 ligand-binding residues. HHsearch–*firestar* predicts 15 078 residues and combining alignments from the two methods helps *firestar* to predict 17 027 ligand-binding residues.

Biological relevance

The PDB contains a diverse range of functional information that can be automatically collected. Unfortunately, much of it is redundant and many PDB files contain artefacts and molecular data without strict biological meaning. Solvent molecules and crystal packing effects produce interatomic contacts between amino acids and heteroatoms that may not have biological relevance. Many PDB structures are crystallized with inhibitors.

FireDB collects and organizes all protein–small ligand interactions in the PDB. FireDB is built around templates generated from a 97% redundant version of the PDB and all protein–small ligand interactions are mapped onto these templates. All functional residues in the FireDB repository are now classified in terms of their biological relevance using evolutionary information, structural data and lists of known cognate ligands. All protein–ligand interactions in FireDB are classified as biologically relevant, putative or non-relevant.

Cognate ligands are those found in PROCOGNATE (19). However, we filter out those that are commonly added as a part of the crystallization process. This excludes most ions, water and solvent molecules such as glycerol. Inhibitors are accepted as ligands as long as they can act as analogues of the cognate ligands.

Evolutionary information for biological relevance analysis is obtained by running *firestar* for all FireDB templates against the FireDB template database. This allows us to cluster together all binding sites that are evolutionarily related. This information is accessible through the FireDB web services and a detailed description is provided in the online help. FireDB also computes the average number of residues that bind each ligand. It has been previously reported that high connectivity is a good descriptor of biological relevance (20).

Biologically relevant protein–ligand interactions are those that involve cognate ligands with at least one evolutionarily related site in FireDB. In the absence of evolutionary information, protein–ligand interactions are considered putative if the ligand is in the cognate list and the number of residues implicated in binding is over two-thirds of the average number for the ligand. Predictions by *firestar* are only made from biologically relevant or putative binding sites.

Further information on the decision-making process involved in determining biological relevance can be found on the web pages.

FireDB is regularly updated with new structures. The greater the amount of functional information in FireDB, the more sequence space can be covered by *firestar*. There has been an increase of 8608 templates with functional sites, since *firestar* was first presented in 2007. The most recent version of FireDB contains 18 048 templates, of which 14 770 contain putative or biologically relevant sites.

The number of binding sites in the database has more than doubled from 38 865 to 86 379, and half of these (41 063) are classified as putative or biologically relevant sites. Only biologically relevant and putative sites are considered by *firestar* for the predictions on the summary pages. The remaining sites are still available through PDB code queries in the FireDB web pages.

High-throughput mode

The *firestar* server has now been enabled to work in high-throughput mode and can be easily integrated into servers either through the server or as a web service. At present it plays an important role as a part of the APPRIS pipeline to annotate splice variants (21) as a part of the ENCODE project. Predictions for the human genome are accessible through APPRIS (appris.bioinfo.cnio.es). The web service differs from the web server in that it predicts only ligand-binding residues and a confidence score for each residue.

FUTURE IMPROVEMENTS

During the CASP9 prediction edition, our group participated with two predictors, the fully automatic server and a version of *firestar* that used 3D models to extend *firestar* predictions. The preliminary results that suggested a slight improvement is obtained by using models.

Given that structural information frequently gives insights about binding mechanisms and ligand-binding specificities, we are working to implement 3D model prediction in *firestar*. Future versions of *firestar* will allow users to retrieve models with the predicted ligand bound to the structure. This is an important feature in which potential users of *firestar* will be interested, even if the improvement in the accuracy of ligand-binding prediction is not always substantial.

We would like to add more sources of annotated functional residues beyond those that are in the PDB and CSA, such as the annotated functionally important residues that are available in a number of sequence databases. Adding further search and alignment methods ought to generate incremental improvements in coverage, although this would affect the performance of *firestar*.

We are working to refine our definition of biological ligands by highlighting those non-cognate ligands that are of pharmacological or chemical importance.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We would like to thank the referees for helping us to improve the *firestar* server and Angel Carro and Jose Maria Fernandez for providing timely technical assistance.

FUNDING

This work was supported by a grant from the Spanish Ministry of Science and Innovation (MICINN, BIO2007-66855) and technical assistance was provided by the Spanish Bioinformatics Institute (INB), a platform of the ISCII. Funding for open access charge: Spanish National Cancer Research Center (CNIO).

Conflict of interest statement. None declared.

REFERENCES

1. Sterk,P., Kulikova,T., Kersey,P. and Apweiler,R. (2007) The EMBL Nucleotide Sequence and Genome Reviews Databases. *Methods Mol. Biol.*, **406**, 1–21.
2. Levitt,M. (2007) Growth of novel protein structural data. *Proc. Natl Acad. Sci. USA.*, **104**, 3183–3188.
3. Lopez,G., Valencia,A. and Tress,M. (2007) FireDB—a database of functionally important residues from proteins of known structure. *Nucleic Acids Res.*, **35**, D219–D223.
4. Lopez,G., Valencia,A. and Tress,M.L. (2007) *firestar*—prediction of functionally important residues using structural templates and alignment reliability. *Nucleic Acids Res.*, **35**, W573–W577.
5. Berman,H.M., Battistuz,T., Bhat,T.N., Bluhm,W.F., Bourne,P.E., Burkhardt,K., Feng,Z., Gilliland,G.L., Iype,L., Jain,S. *et al.* (2002) The Protein Data Bank. *Acta. Crystallogr. D Biol. Crystallogr.*, **58**, 899–907.
6. Wang,C., Vernon,R., Lange,O., Tyka,M. and Baker,D. (2010) Prediction of structures of zinc-binding proteins through explicit modeling of metal coordination geometry. *Protein Sci.*, **19**, 494–506.
7. Wass,M.N., Kelley,L.A. and Sternberg,M.J. (2010) 3DLigandSite: predicting ligand-binding sites using similar structures. *Nucleic Acids Res.*, **38**, W469–W473.
8. Fischer,J.D., Mayer,C.E. and Soding,J. (2008) Prediction of protein functional residues from sequence by probability density estimation. *Bioinformatics*, **24**, 613–620.
9. Lopez,G., Rojas,A., Tress,M. and Valencia,A. (2007) Assessment of predictions submitted for the CASP7 function prediction category. *Proteins*, **69(Suppl. 8)**, 165–174.
10. Lopez,G., Ezkurdia,I. and Tress,M.L. (2009) Assessment of ligand binding residue predictions in CASP8. *Proteins*, **77(Suppl. 9)**, 138–146.
11. Porter,C.T., Bartlett,G.J. and Thornton,J.M. (2004) The Catalytic Site Atlas: a resource of catalytic sites and residues identified in enzymes using structural data. *Nucleic Acids Res.*, **32**, D129–D133.
12. Altschul,S.F., Madden,T.L., Schaffer,A.A., Zhang,J., Zhang,Z., Miller,W. and Lipman,D.J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.*, **25**, 3389–3402.
13. Soding,J. (2005) Protein homology detection by HMM-HMM comparison. *Bioinformatics*, **21**, 951–960.
14. Tress,M.L., Grana,O. and Valencia,A. (2004) SQUARE—determining reliable regions in sequence alignments. *Bioinformatics*, **20**, 974–975.
15. Zemla,A. (2003) LGA: a method for finding 3D similarities in protein structures. *Nucleic Acids Res.*, **31**, 3370–3374.
16. Zhang,Y. (2008) I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics*, **9**, 40.
17. Tress,M.L., Jones,D. and Valencia,A. (2003) Predicting reliable regions in protein alignments from sequence profiles. *J. Mol. Biol.*, **330**, 705–718.
18. Harrow,J., Denoeud,F., Frankish,A., Reymond,A., Chen,C.K., Chrast,J., Lagarde,J., Gilbert,J.G., Storey,R., Swarbreck,D. *et al.* (2006) GENCODE: producing a reference annotation for ENCODE. *Genome Biol.*, **7(Suppl. 1)**, S4.1–9.
19. Bashton,M., Nobeli,I. and Thornton,J.M. (2008) PROCOGNATE: a cognate ligand domain mapping for enzymes. *Nucleic Acids Res.*, **36**, D618–D622.
20. Dessimailly,B.H., Lensink,M.F., Orengo,C.A. and Wodak,S.J. (2008) LigASite—a database of biologically relevant binding sites in proteins with known apo-structures. *Nucleic Acids Res.*, **36**, D667–D673.
21. Tress,M.L., Wesselink,J.J., Frankish,A., Lopez,G., Goldman,N., Loytynoja,A., Massingham,T., Pardi,F., Whelan,S., Harrow,J. *et al.* (2008) Determination and validation of principal gene products. *Bioinformatics*, **24**, 11–17.